The Thesis Committee for Erika Signa Benson-Villegas
Certifies that this is the approved version of the following thesis:

Lexical Revisions and Filled Pauses: Associations with Vocabulary Knowledge in Bilingual Children

APPROVED BY
SUPERVISING COMMITTEE:

Supervisor: _____
Lisa M. Bedore

_____
Courtney T. Byrd

Lexical Revisions and Filled Pauses: Associations with Vocabulary Knowledge in Bilingual Children

by

Erika Signa Benson-Villegas, B.A.

Thesis

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Arts

The University of Texas at Austin

May 2015

Acknowledgements

I would like to express my wholehearted gratitude to my advisor, Lisa M. Bedore, whose enthusiasm, knowledge, and dedication has been invaluable throughout my research and graduate school career. I would also like to express sincere thanks to Courtney T. Byrd for her thoughtful feedback, insights, and genuine encouragement. I am profoundly grateful to the Human Abilities in Bilingual Language Acquisition (HABLA) lab team, who were instrumental in not only data collection and navigating data files but for fostering a welcoming and supportive environment. I would like to express special thanks also to John and Gloria Benson, Ingrid Benson, and Adam Levy. Their love and support has been a strong, motivating force that has kept me looking forward.

Abstract

Lexical Revisions and Filled Pauses: Associations with Vocabulary Knowledge in Bilingual Children

Erika Signa Benson-Villegas, M.A.

The University of Texas at Austin, 2015

Supervisor: Lisa M. Bedore

Purpose: The current study examined lexical revisions and filled pauses as related to vocabulary knowledge and language experience in bilingual children with and without language impairment.

Method: Participants included Spanish-English speaking bilingual children (n=30) aged 84-103 months. Children were designated into one of three language groups based on language ability (typically-developing, language-impaired, and at-risk typically-developing). Narratives from the Test of Narrative Language were transcribed and coded in Spanish and English for lexical revisions and filled pauses. Lexical revisions and filled pauses in each language were then correlated with measures of vocabulary knowledge provided through the narrative sample (MLU and NDW) and standardized testing (EOWPVT and BESAME Semantics scores).

Results: The current study found that the typically-developing group exhibited significantly more lexical revisions in Spanish and significantly more filled pauses in

English and Spanish compared to the at-risk typically-developing and language-impaired groups. NDW significantly correlated with total maze use, lexical revisions, and filled pauses in both English and Spanish. No significant correlations were observed between language dominance or language exposure to either lexical revisions or filled pauses. Lexical revisions and filled pauses were significantly correlated to each other within each language, and significantly correlated across the languages.

Conclusions: Results suggest that lexical revisions and filled pauses are more closely related to vocabulary knowledge than to language dominance or exposure. Lexical revisions and filled pauses were produced more in children with higher levels of vocabulary knowledge as measured by NDW as opposed to MLU or standardized test measures. Lexical revisions and filled pauses demonstrated correlations across and within languages, indicating that language experience does not seem to play a role in their use.

Table of Contents

List of Tables

List of Figures

INTRODUCTION

MAZES

Mazes may be a venue for understanding the metalinguistic and cognitive demands of language formulation. Loban (1976) defined mazes (also referred to as language disfluencies, language revisions, or language repairs) as a series of words, initial parts of words, or unattached fragments which do not contribute meaning to the ongoing flow of language. Mazes are a normal part of speech production, yet they have significance because their frequency, structure, and distribution can help to distinguish language disorders (Nettelbladt & Hansson, 1999). Current research suggests that these language disfluencies are caused by either decreased linguistic knowledge or increased language awareness. An increase in language demands will tend to further challenge language abilities and awareness, thereby causing mazes.

Mazes have been noted to be prevalent in the language of children with language impairment (LI). Bilingual children have also been noted to have an increased use of mazes when compared to monolingual children. Children with language impairments have been observed to talk less than their typically-developing peers and the occurrence of mazes in their language is more prevalent (MacLachlan & Chapman, 1988; Dollaghan & Campbell, 1992; German, 1994; Miller & Klee, 1995, as cited in Nettelbladt & Hansson, 1999).

Levelt (1989) proposed that mazes are caused by linguistic uncertainty resulting from the demands of language production. For instance, when a speaker is unable to retrieve or produce a word, uncertainty may manifest at the grammatical, phonological,

or lexical level (Levelt, 1989). These moments of uncertainty may cause the speaker to detect linguistic breakdowns that require repairs manifested as mazes. Particularly, vocabulary knowledge is decreased in children with language impairment, and a high use of mazes may correspond to difficulty with retrieving a word, especially as uncertainty in word retrieval is tied to maze production (Bedore et al., 2006). Mazes may also provide insight into underlying metalinguistic and cognitive processes involved in language planning and formulation. The information gained by researching mazes includes knowledge of a child's planning, memory, and stage of language acquisition, factors which are significant for diagnosis and treatment.

BILINGUALISM AND MAZE USE

Bilingual speakers tend to produce more mazes in their second language than in their first language (Wiese, 1984; Lennon, 1990; Poulisse, 1999; Rieger, 2003). The current question in regards to bilingual maze use is discerning whether mazes are produced by increased linguistic uncertainty or by decreased language knowledge. Bilinguals tend to have more language knowledge due to knowing two languages yet may have more linguistic uncertainty as a result of using each of their languages less than their monolingual peers (Bedore et al., 2006). Kormos (1999) reports that the ability to revise in a second language requires knowledge of the second language and does not only reflect uncertainty. Bilinguals purportedly have more tip of tongue phenomena, which has been attributed to differences in the strength of connections between semantic and phonological information in bilinguals (Gollan & Silverberg, 2001; Gollan & Acenas,

2004; Gollan, Montoya & Bonanni, 2005). Tip of the tongue phenomena may be associated with increased linguistic uncertainty (Bedore et al., 2006).

Maze use also changes over time as fluency in L2 increases. In a study by Verhoeven (1989), Turkish children's L2 proficiency in Dutch was associated with a shift from phonological revisions to morphosyntactic revisions. Poulisse (1997) purported that children in a study by Wiese (1984) may have produced more repetitions, corrections, and filled pauses in L2 than L1 due to lack of automaticity in L2. Revisions may also be different in each language as a result of syntactic structure and item differences. For example, Rieger (2003) found that German-English bilinguals revised different elements in each language. In English, speakers repeated personal pronouns and prepositions; however, in German, speakers most commonly repeated demonstrative pronouns. The pattern in Rieger (2003) was attributed to increased preposition use in English relative to German, thereby creating more opportunities for revising prepositions in English than German. Bedore et al. (2006) found that in Spanish narrative samples, grammatical revisions were more prevalent than in English. Bedore et al. (2006) also noted that grammatical revisions were positive correlated with the number of different words in Spanish but in English the number of different words correlated to the use of filled pauses. These findings suggest that speakers with larger vocabularies may have more uncertainty in word selection but the different languages cause the uncertainty to be manifested in different types of mazes.

LANGUAGE MONITORING

Creating and producing a message relies on metalinguistic abilities. One such theory of language formulation is the perceptual loop theory, composed by Levelt (1989). The perceptual loop theory is based on spreading activation and self-monitoring of internal and overt speech. The first component is conceptual preparation, in which word generation moves through lexical selection, morphological and phonological encoding, phonetic encoding, and articulation. Meanwhile, the speaker is monitoring self-produced internal and overt speech (Levelt, Roelofs, & Meyer, 1999). In perceptual loop theory, three loops work sequentially to monitor the language output. The first loop occurs when the speaker formulates a message, which is then compared to the speaker's original intentions. The second loop is known as covert or pre-articulatory monitoring and occurs when the message is monitored before articulated. The third loop of monitoring then occurs after the utterance is articulated.

In contrast to the perceptual loop theory, bilingual language production models take into account how speech production is affected by the presence of two languages. Bilinguals have been reported to be more proficient on non-linguistic tasks requiring cognitive control; yet seem relatively less proficient at rapid lexical access and tasks that require vocabulary knowledge. Two such bilingual language production models are the frequency-lag hypothesis and the competition for production model. The frequency-lag hypothesis assumes that bilinguals have fewer opportunities to speak each of their languages. Since each language is used less compared to monolinguals, production may be slower in both of their languages because there is reduced accessibility, or lower

frequency, of words in their lexicon (Gollan, Montoya, Cera, & Sandoval, 2008; Gollan, Montoya, Fennema-Notestine, & Morris, 2005; Gollan et al., 2011). The frequency-lag hypothesis is observed in the difference in recall between low-frequency and high-frequency words, evidenced by bilinguals having more difficulty recalling low-frequency versus high-frequency words. In addition to lexical access, the frequency-lag hypothesis may also reflect interactions across two languages since bilingualism requires different lexical mappings and grammatical commitments than monolingualism. Although the frequency-lag hypothesis explains the bilingual disadvantage seen in lexical processing, the hypothesis does not explain how bilingual language production causes cognitive advantages. In contrast, the competition for production model provides an explanation for how inhibition of L1 may engage cognitive mechanisms that lead to the cognitive advantages seen in bilinguals (Kroll & Gollan, 2014).

The competition for production model assumes that both languages are activated when speech is planned, which requires that the languages compete before a target language is designated (Kroll & Gollan, 2014). Bilinguals are reported to have more Tip of Tongue (ToT) states, which have been suggested to occur due to an increased generalized load of their lexical system. However, in a study by Gollan et al. (2005), bilinguals were as successful as monolinguals at retrieving proper names, indicating that competition between languages is not the only cause of more TOT states in bilinguals. In an ERP study, Hoshino and Thierry (2011) found evidence that cross-language activation occurs relatively early in speech planning. Kroll and Gollan (2014) reviewed imaging and behavioral studies and concluded that L1 appears to be active even without being

engaged. In addition, Kroll and Gollan (2014) noted that L2 activation causes dramatic interference in L2 production. L1 activation during L2 speech planning leads to inhibition of dominant L1, per the inhibitory control model (Green, 1998). The competition for production model depicts how bilingual planning produces cognitive effects during language choice and activation as well as how the activation of both languages may affect lexical access. Both bilingual speech production models, frequency-lag and competition for production, are based on the interactions of two languages which ultimately affect lexical access and language change in a bilingual individual over time (Kroll & Gollan, 2014).

THE ROLE OF COGNITIVE PROCESSES IN LANGUAGE MONITORING

Two cognitive processes involved in language learning and language monitoring are attention and working memory. Working memory is affected by the demands of a task and can influence to what extent speakers are capable of monitoring their speech (Harrington & Sawyer, 1992). For example, the task of narrative story-telling is syntactically demanding and involves both composition, synthesis, and delivery of complex information in a complex format (Taliancich-Klinger, Byrd, & Bedore, 2013). Narratives also allow the speaker to have more control over word choice and grammatical structures, which may contribute to increased linguistic uncertainty in the narrative context (Bedore et al., 2006). Thus, if attention and working memory are devoted to planning and other aspects of the narrative task, then language formulation and

production may be monitored to a lesser extent due to those resources being dedicated to other causes.

When there are two languages in a language system, monitoring may occur differently for each because the dominance patterns of the two languages may vary. A less dominant language would require more attention and monitoring than a more dominant, automatic language. For example, as emerging bilinguals learn a second language (L2), syntactic rules and vocabulary are less automatic than syntactic rules and vocabulary in their dominant first language (L1). Consequently, the L2 might require more attention and monitoring and exhibit more prevalent maze use.

Mazes may be more evident in L2 learning (Robinson, 1995). For example, when an L2 speaker is uncertain whether an utterance contains an error, monitoring allows the speaker to notice the gap in one's knowledge and trigger further acquisition processes (Robinson, 1995, as cited in Kormos, 1999). L2 monitoring involves the checking of both internal and external speech against the existing linguistic system. L2 monitoring is assumed to be similar to comprehension and relies on receptive knowledge (de Bot, 1996, p. 551). Robinson (1995) stated that noticing gaps in L2 knowledge involves the speaker's ability to detect and rehearse in short-term memory, and then encode in long-term memory. Noticing errors henceforth advances L2 learning, and rehearsal assists in encoding the correct format in long-term memory.

TYPES OF MAZES IN LANGUAGE-IMPAIRED CHILDREN

Mazes can be categorized as filled pauses, repetitions, and revisions (DeJoy &Gregory, 1985; Dollaghan & Campbell, 1992). Revisions can be classified further into grammatical revisions, lexical revisions, and phonological revisions (Dollaghan & Campbell, 1992). Repetitions also can be classified further into partial, word, or phrase repetitions (Bedore et al., 2006). Van Hest (1996b) and Poulisse and Bongaerts (1994) found that content words were more often corrected than function words due to content words carrying more information. German (1994) mentioned mazes as part of the typical behavior of a subgroup of children with word-finding problems, and recommended the use of narratives to study the behavior.

Nettelbladt and Hansson (1999) found that children with SLI differed from children with phonological impairment in having more mazes, more repetitions, more pauses, repeating initial phonemes and syllables rather than whole words, and repeating lexical words as often as function words. Similarly, Thordardottir and Ellis Weismer (2002) compared the language in narratives of children between the ages of 5 and 9 with and without LI and found that children with LI had more prevalent content mazes, or lexical and syntactic revisions, and that mazes were produced more often in children with LI in general.

Previous research has demonstrated that in both LI and control groups of phonological impairment or typically-developing language, MLU is higher in utterances containing mazes than the total MLU (Miller, 1987; MacLachlan & Chapman, 1988; and Netterbladt & Hansson, 1999). DeJoy and Gregory (1973) also found that 4-year-old

children demonstrated a significantly higher percentage of disfluencies with longer sentences (as cited in Gordon, Luper, & Peterson, 1986).

Researchers have interpreted filled pauses and repetitions as indicating immature disfluencies that are expected to decline with increased language development (DeJoy & Gregory, 1985; Starkweather, 1987). Starkweather (1987) reported that as language develops in children, their speech becomes more fluent, thereby exhibiting less maze use. Loban (1976) also reported that individuals with a high language proficiency exhibited less maze behavior. However, Kaur, Hegde, Kumaraswamy, and Rao (2011) found that children who were less fluent in English had less ability to rectify and manipulate their mazes. When speakers demonstrate difficulties formulating language and still do not use self-repairs, they may be showing that they have more difficulty processing complex linguistic information (Fletcher, 1991; Restrepo, 1996). For instance, children with language impairment (LI) have also been noted to make less use of self-repairs and produce a greater number of non-detected errors (Navarro-Ruiz & Rallo-Fabra, 2001). Less use of self-repairs suggests that their rate of revisions would be lower than a typically-developing child. Yet, children with LI have been noted to produce more mazes than children without LI (Leadholm & Miller, 1995, as cited in Bedore et al., 2006).

VOCABULARY

When compared to monolingual children, bilingual children show less vocabulary knowledge if only one of their languages is examined. Yet, the lexical repertoire of the bilingual child actually contains about the same number of words at similar points in

development as monolingual children. The lexical system of bilingual children is also acquired and organized similarly (Holowka et al., 2002; Patterson, 1998, 2000; Pearson et al., 1993, 1995). The difference between monolingual and bilingual vocabulary knowledge lies predominantly in the way bilingual vocabulary knowledge overlaps across the two languages (Pearson & Fernandez, 1994; Peña et al., 2002).

A bilingual child's vocabulary knowledge is influenced by the amount of exposure received in each language (Marchman et al., 2004; Patterson, 2000). In a study by Peña et al. (2002), bilingual Spanish-English speaking children named similar numbers of words in each language in a category generation task; interestingly, they tended to generate different items in each language, which correlated to specific activities in specific language contexts.

MLU in words is a measure for assessing vocabulary knowledge. Paradis et al. (2003) found that bilingual children tended to have higher MLU and greater vocabulary knowledge in their dominant languages. Children with higher MLUs have also been found to produce more grammatical and lexical revisions, indicating that their metalinguistic knowledge is sufficient to make these revisions (Bedore & Peña, 2008). A measure of high MLU therefore suggests that grammatical and lexical revisions are expected to be more prevalent in children who have greater command of their language skills.

Semantic knowledge is significant as a potential clinical factor in diagnosis because difficulties with word retrieval, word meaning, and word learning would be comparable in both languages if a child is language-impaired (Bedore & Peña, 2008). For

example, Sheng, Peña, Bedore, and Fiesta (2012) found that bilingual children with LI have sparsely linked semantic networks comparable to monolingual children with LI. In addition, a study by Peña, Iglesias, and Lidz (2001) found that typically-developing bilingual children made significant gains in a single-word labeling task intervention, whereas bilingual children with low language abilities made minimal to no gains. Gutiérrez-Clellen and DeCurtis (1999) compared the quality of definitions produced by bilingual children with and without LI and found that children with LI used nonspecific vocabulary, provided infrequent elaborations, and were unable to account for multiple or colloquial word meanings. Fiestas, Peña, Bedore, and Sheng (2011) also found that Spanish-English bilingual children with LI had difficulties describing functions of object nouns and providing category labels.

Bilingual children with LI with low vocabulary knowledge may exhibit either less or more lexical revisions and filled pauses than TD children with a higher vocabulary knowledge due to decreased linguistic uncertainty (causing more lexical revisions and/or filled pauses) or less metalinguistic ability to self-repair (causing less lexical revisions and/or filled pauses).

THE CURRENT STUDY

The current study was designed to examine the role of lexical revisions and filled pauses in specific relation to vocabulary knowledge. We will evaluate lexical revisions and filled pauses exhibited in narrative language samples elicited by the *Test of Narrative Language (TNL)* and vocabulary knowledge as tested in the *Expressive One-Word Picture*

Vocabulary Test (EOWPVT) and *Bilingual English-Spanish Assessment Semantics subtest*. Through narrative language sample analysis using the *Systematic Analysis of Language Transcripts (SALT), Expressive One-Word Picture Vocabulary Test (EOWPVT),* and *BESAME Semantics* scores of bilingual children who are typically-developing (TD), language-impaired (LI), and at-risk typically-developing, the following questions will be addressed:

Do patterns of lexical revision and filled pauses in children with typically developing language skills differ by language ability?

Is vocabulary knowledge correlated with higher use of either lexical revisions or filled pauses?

Is MLU correlated with higher use of either lexical revisions or filled pauses?

Do bilingual children with more dominance and exposure to one language demonstrate differences in lexical revisions and filled pauses in each of their languages? If so, to what degree and will either lexical revisions or filled pauses be more affected?

METHOD

PARTICIPANTS

Thirty English-Spanish bilingual participants were selected from a larger sample of children recruited for a longitudinal research study. The participants were divided into three groups of ten according to whether their test scores indicated them as having typically-developing language (TD), language-impairment (LI), or low-normal typically-developing language (LN). All participants were 2nd graders at the time of testing. The mean age of the participants was 91.6 months old, or about 7.58 years of age. All participants were recruited from the local Austin-area school districts of Hays ISD and Pflugerville ISD.

Children were selected to participate in the longitudinal study if they had at least 20% input and output in Spanish and English across home and school as determined by language profile questionnaires filled out by the parent and information gathered during a parent telephone interview. Children were excluded if they had hearing loss, severe social-emotional problems, mental retardation, autism spectrum disorder, a brain injury, or low cognitive ability, indicated by scoring below the cut-off score of 75 on the *Universal Non-Verbal Intelligence Test* (Bracken & McCallum, 1998).

RECRUITMENT MEASURES

Children were assigned a number for tracking, privacy, and confidentiality. Verbal and/or written assent was obtained from participants in the study if they were 6 or

older. Children were tested in both English and Spanish in a quiet room in their elementary school. Testing in each language occurred on different days. Children were tested by trained bilingual research assistants or graduate students from the University of Texas at Austin's Communication Sciences and Disorders program. Examiners attempted for the children to adhere to the target language. Children's responses were recorded via digital recorders and then transcribed for scoring.

Phase 1. Phase 1 was organized as a screener to determine the participants' language profiles, assess language dominance, and screen for language impairment. Participants were contacted through their schools. Families provided active written consent if they were interested in participating and if they felt they met the study's inclusionary criteria. The principal components of Phase I were administering the *Bilingual English Spanish Oral Screener (BESOS)* and conducting a language experience interview with the parents.

BESOS. The BESOS was used to screen for language impairment, typically developing language, and at-risk typically developing language (low-norm). All children with comparable language experience were invited to participate. The composite score of the BESOS, which includes both semantics and syntax scores in both languages, indicated into which of the three groups the participants would fall. If a participant scored below the 15th percentile, they scored within the LI range, thereby joining the LI group. If the participant scored within the 15th to 30th percentile ranges, they were

classified into the at-risk typically-developing language group. Finally, if a participant scored over the 30th percentile, they fell into the TD group.

Parent Language Interview. Language profile and socioeconomic information was obtained from a telephone interview with the participants' parents. Parents provided their child's age of exposure to the English language and their own education and occupation levels. A socioeconomic status (SES) value was then computed from the education and occupation levels provided. The parent interview also investigated the amount of time children heard and used Spanish and English on weekdays and weekends.

Phase 2. Phase 2 was conducted one year after Phase 1. Phase 2 consisted of developmental testing in English and Spanish, which were then used to classify children into one of three language groups: typically-developing, at-risk typically developing language (low-norm), and language-impaired. Developmental testing consisted of: *Test of Narrative Language (TNL), Bilingual English Spanish Assessment (BESA), Expressive One-Word Picture Vocabulary Test (EOWPVT), Universal Nonverbal Intelligence Test (UNIT), Non-Word Repetition,* and parent and teacher interviews.

Test of Narrative Language (TNL). The TNL is an assessment that measures the ability of children aged 5;0 -11;11 to comprehend and convey narratives. The *TNL* identifies language impairments, measures the ability to answer literal and inferential comprehension questions, and measures how well children use language in narrative discourse. There are three narrative formats in the *TNL*. The first narrative format is a story retell with no visual cues. The second narrative format is a story formulation task

with five sequenced pictures, which the participant tells after listening and answering comprehension questions about a similar story. The third narrative format is another story formulation task which the child narrates using one picture.

The Spanish and English *TNL*s have similar structures. The English *TNL* is a norm-referenced test that provides standard scores, percentile ranks for Narrative Comprehension and Oral Narration, and an overall standard score called the Index of Narrative Language Ability. The Spanish *TNL* is an experimental version that was adapted from the English *TNL* which has been shown to reliably differentiate between bilingual children with and without language impairment. The stories in the English and Spanish *TNL*s are different and not direct translations of each other. Participants were tested in the standardized English *TNL* and the experimental Spanish *TNL* for each of the three narrative formats.

UNIT. *UNIT* scores were used to evaluate participants' nonverbal problem solving skills and establish language group placements. The *UNIT* is administered non-verbally and provides an assessment of intelligence for children from 5 through 17 years of age. If participants scored below a cut-off of 75, then they were excluded from the study.

BESAME. The *BESAME* was administered in both English and Spanish to children aged 7-9;11. It was developed following the developmental patterns of each language. Semantic subtests use conceptual scoring that is effective for bilingual children (Bedore, Peña, Garcia, & Cortez, 2005). Morphosyntax subtests focus on structures that

are difficult for children with SLI (Bedore & Leonard, 2001). Based on the responses given to the *BESAME*, a child's language development could be analyzed for changes in semantics and syntax domains.

EOWPVT. The *Expressive One-Word Picture Vocabulary Test* (2000) and the *Expressive One-Word Picture Vocabulary Test-Spanish Bilingual Edition* (2001) are individually administered, norm-referenced tests of single-word expressive vocabulary. The *EOWPVT* required children to name pictures in English and Spanish to determine the level of their vocabulary knowledge. 190 items are presented in developmental sequence, which are included in both test editions. Basal and ceiling rules specific to *EOWPVT* were followed to score both the English and Spanish Editions. Test administration was discontinued if a basal was not reached. After the ceiling was reached, 14 additional items were administered to ensure a ceiling was reached across all scoring methods. The English version of the *EOWPVT* contains 20 items that are not administered in the standardized administration of the *EOWPVT-SBE*. To compare item level results in both tests, administration procedures were modified for the bilingual *EOWPVT-SBE*. The *EOWPVT-SBE* was conducted in Spanish, and responses were elicited in Spanish. English and bilingual versions were administered as English only and Spanish only versions. All responses were recorded verbatim during both administrations. Responses were elicited in the target language, even if they were first given in another language.

Parent and Teacher Interviews: Parent and teacher interviews were used to document children's use and exposure to Spanish and English. Caregivers and teachers

independently responded to a questionnaire. The questionnaire asked about the participants' hour-by-hour exposure and use of Spanish and English at home and in the classroom. They were also asked to rate the participants' ability in: frequency of language use with peers and adults, vocabulary, speech, sentence production, grammatical, and comprehension proficiency (Gutiérrez-Clellen & Kreiter, 2003).

CLASSIFICATION

The children were classified into one of three language groups (typically-developing (TD), language-impaired (LI), and low-normal (LN)) based on converging evidence from parent/teacher language concern, the results of the *BESOS*, their best *BESA* semantics and morphosyntax score, and their *TNL* scores. A formula leading to a numerical value indicated whether they were TD (0-1.5), LN (2-3.5), or LI (4+).

PROCEDURES

Participant Selection. The current study selected thirty children from the longitudinal study to compose three language groups (typically-developing, low-normal, and language-impairment) of ten children each. All children in the three groups were second graders. They were comparable in age and socioeconomic status (SES). Age ranged from 89-99 months in the TD group, 90-103 months in the LN group, and 84-101 months in the LI group. The mean ages were 94.8 for the TD group, 95.5 for the LN group, and 92.1

for the LI group. The means for SES across all groups were fairly close, with the TD and LN SES calculated at 2.2 and the LI group calculated at 2.00. The language exposure was above 50% in Spanish for TD and LI groups at 59.941% and 63.709% respectively and slightly below 50% in Spanish for the LN group at 48.185%. Gender was evenly matched for each of the groups. The TD group consisted of four males and five females; the LN and LI groups consisted of five males and five females each. Narratives were transcribed for all children in the study in English and Spanish. If the sample was not considered sufficient or representative of their language, then the participant was excluded. *BESAME Semantics* and *EOWPVT* scores in Spanish and English were also required for participation in the study. Table 1 depicts the age, SES, age of English exposure, percent of combined input and output in each language, and gender per group.

Table 1. Demographic Features of Participants

Mean Ages, SES, Age of English Exposure, Percent of English/Spanish Exposure, & # of Males/Females

Classification	Age	SES	AoE exposure	Males	Females	% English Input Output	% Spanish Input Output
Typically-Developing	94.8 (89-99 range)	2.20	3.278	4	6	40.059%	59.941%
Low-Normal	95.5 (90-103 range)	2.20	2.1	5	5	51.815%	48.185%
Language-Impaired	92.1 (84-101 range)	2.00	4.5	5	5	36.291%	63.709%

Scoring.

TNL. Narratives from the TNL were transcribed by trained bilingual graduate and undergraduate students using Systematic Analysis of Language Transcripts (SALT) program and conventions. Utterances were segmented into communication units (C-units), words and morphemes were coded, and mazes were marked within the participants' utterances. Intelligible utterances were included in the analysis. Unintelligible utterances were excluded. Guidelines for spoken narrative production outlined by Loban (1976) were followed. Mazes were coded for filled pauses (FP), repetitions (REP), connectors (CON), grammatical revisions (GREV), lexical revisions (LREV), and phonological revisions (PREV). Narratives were also analyzed for total

mazes produced, Mean Lengths of Utterance (MLU), and Number of Different Words (NDW).

EOWPVT. Five scores were computed from the *EOWPVT* and *EOWPVT-SBE*: *Monolingual English, Monolingual Spanish, Within-Test Conceptual Scores SBE Edition,* and *Across Test Conceptual Scores* in English and Spanish. Monolingual English standard scores were derived using English norms. Monolingual Spanish standard scores were derived using norms in the *EOWPVT-SBE* manual. To be accepted as correct in Monolingual English and Spanish scoring, responses must have been given in the target language. For Within Test Conceptual Scores (SBE) and Across Test Conceptual Scores, responses in both languages were accepted as correct. Standard scores for Within Test Conceptual Scores SBE were derived using norms in the SBE manual. Across Test Conceptual Scores were derived by analyzing each item across both administrations. Two standard scores were computed from Across Test Conceptual Scores. One score was derived from the SBE manual and the second score derived from the English edition. Tables 2 and 3 illustrate the means and standard deviations of the BESAME Semantics and EOWPVT raw within-language and single-language scores in Spanish and English.

Table 2.

Spanish BESAME Semantics Scores and EOWPVT Raw Scores

	BESAME SEMANTICS		EOWPVT WITHIN-LANGUAGE		EOWPVT SINGLE-LANGUAGE	
Group	Means	Standard Deviation	Means	Standard Deviation	Means	Standard Deviation
TD	29.9	8.3779	57.1	10.1593	53.0	7.211
LN	22.7	6.1833	47.5	7.487	38.8	17.1969
LI	15.0	5.3955	43.2	10.3580	40.1	8.2118

Table 3.

English BESAME Semantics Scores and EOWPVT Raw Scores

	BESAME SEMANTICS		EOWPVT WITHIN-LANGUAGE		EOWPVT SINGLE-LANGUAGE	
Group	Means	Standard Deviation	Means	Standard Deviation	Means	Standard Deviation
TD	26.3	7.2732	49.5	20.1288	47.8	19.8035
LN	20.7	5.3759	45.4	13.1589	45.4	13.1589
LI	13.6	8.8854	29.4	21.3656	23.2	20.6548

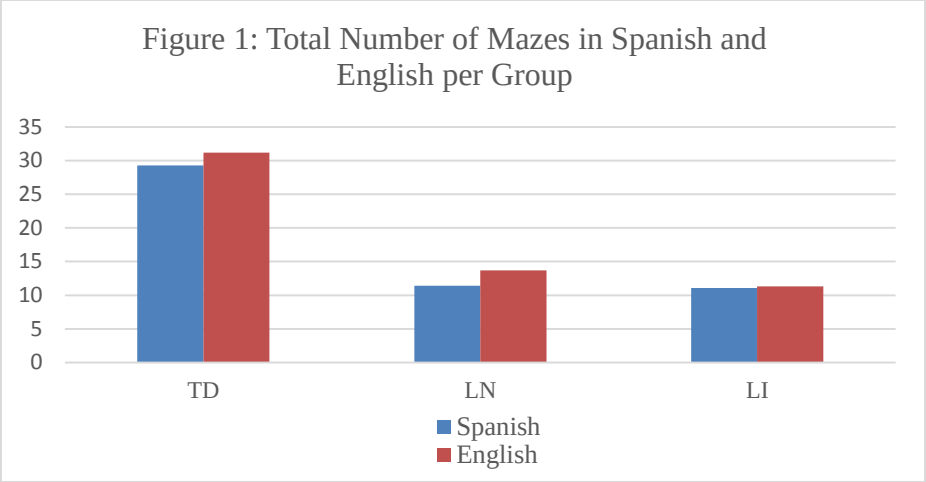
Analyses. All analyses were conducted using SPSS software version 21 (SPSS Inc., Chicago, Illinois, U.S.A.). ANOVAs were used to examine *BESAME Semantics* scores, *EOWPVT* scores, MLU, NDW, total mazes, lexical revisions, filled pauses, and language experience between the three different groups (LI, TD, and LN) in English and Spanish. We then conducted correlations to investigate whether specific standardized measures, NDW, or MLU correlated to maze behavior. Lastly, we examined whether language experience via language exposure and age of acquisition were correlated to types or total number of maze.

RESULTS

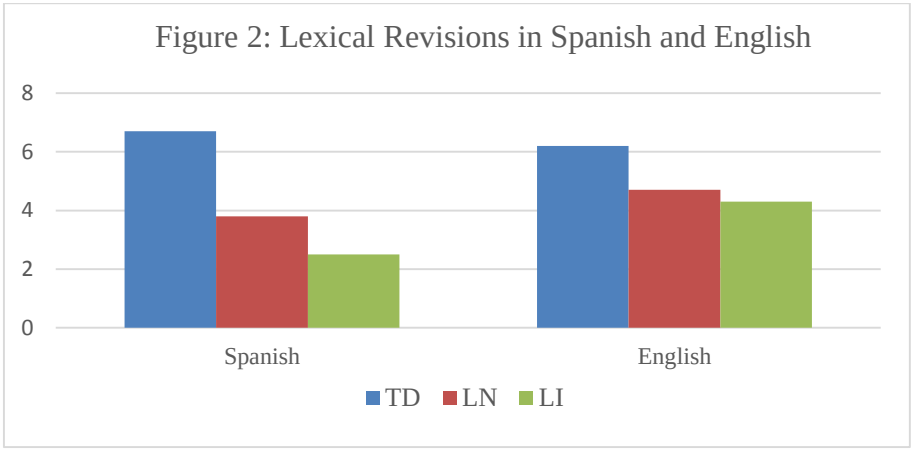
LEXICAL REVISIONS AND FILLED PAUSES BY LANGUAGE ABILITY

The first question addressed whether mazes, lexical revisions, and filled pauses differed by language ability. A correlation analysis found that the number of mazes, lexical revisions, and filled pauses appeared to differ by language ability in both English and Spanish. The typically-developing group presented with more total mazes, lexical revisions, and filled pauses than the low-normal or language-impaired groups. Scheffe's contrast analysis showed that the typically-developing group presented with significantly more filled pauses in English and Spanish and lexical revisions in Spanish than either low-normal or language-impaired groups. Overall, lexical revisions and filled pauses were observed most often in English with the exception that TD children used more lexical revisions in Spanish than in English.

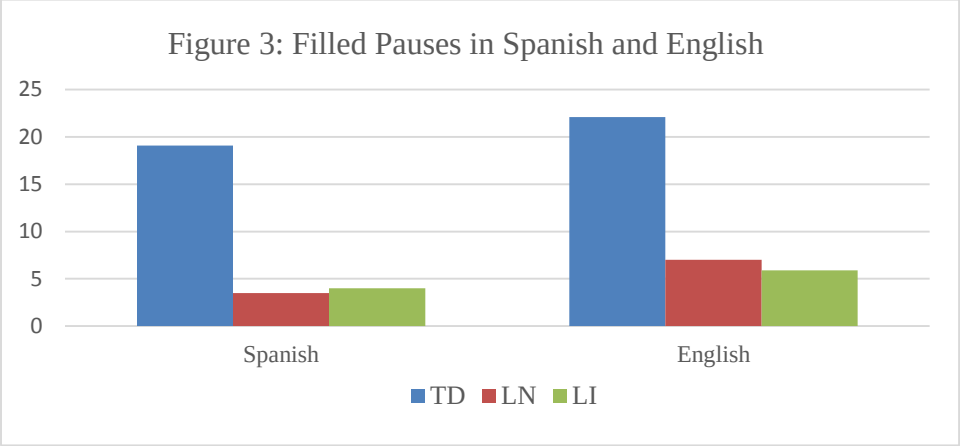
Total Mazes in English and Spanish. In Spanish, children with typically-developing language were found to have a total maze count of 29.3, whereas children with low-normal language had a total maze count of 11.4 and children who were language-impaired had a maze count of 11.1. In English, children with TD language had 31.2 mazes, children with LN language had 13.7, and children with LI had 11.3. These numbers are represented in Figure 1.



Lexical Revisions. For lexical revisions in Spanish, the TD language group used 6.7, the LN language group used 3.8, and LI language group used 2.5. The TD group was significantly different from the LI group per Scheffe's contrast analysis (0.022). In English, the TD group used 6.2, the LN groups used 4.7, and the LI group used 4.3. In a cross-linguistic comparison, lexical revisions were observed more often in English for LN and LI children. Unlike the LN and LI groups, TD children had more lexical revisions in Spanish (6.7) than in English (6.2). Lexical revisions per group in English and Spanish are depicted in Figure 2.



Filled Pauses. In Spanish, the TD language group used 19.1 filled pauses, the LN group used 3.5, and the LI group used 4.00. In English, the TD group used 22.1, the LN group used 7.00, and the LI group used 5.9. TD children had a much higher prevalence of filled pauses than LN or LI groups. Per Scheffe's contrast analysis, the TD group's use of filled pauses significantly differed from both LN and LI groups in English and Spanish. All three groups demonstrate more filled pauses in English than Spanish. Filled pauses are depicted in Figure 3.



STANDARDIZED VOCABULARY MEASURES, LEXICAL REVISIONS, AND FILLED PAUSES

The second question addressed whether vocabulary knowledge was correlated with higher use of either lexical revisions or filled pauses. Vocabulary knowledge was measured by EOWPVT within-language scores, EOWPVT single-language conceptual scores and BESAME Semantics scores. ANOVA results indicated that the Spanish EOWPVT raw within-language scores and BESAME Semantics scores in Spanish and English significantly differentiated between the three language groups by ability when

using Scheffe's contrast analysis; however, the English raw EOWPVT within-language score did not demonstrate differentiation of these three groups.

As language scores increased across TD, LN, and LI groups, so did the total number of mazes. Lexical revisions and filled pauses in Spanish correlated to standardized vocabulary measures; however, the pattern was not reciprocated in English.

Lexical Revisions. In English, lexical revisions were not significantly correlated to any of the vocabulary standardized test measures. In Spanish, lexical revisions were significantly correlated to both EOWPVT within-language scores and single-language conceptual scores.

Filled Pauses. In English, filled pauses were not significantly correlated to any of the standardized vocabulary measures. In Spanish, filled pauses were significantly correlated to BESAME Semantics scores and moderately correlated to both EOWPVT within-language and EOWPVT single-language conceptual scores. Table 4 illustrates the correlations of lexical revisions and filled pauses to vocabulary measures in Spanish while Table 5 illustrates the correlations in English.

Table 4.

Pearson Correlations of Lexical Revisions and Filled Pauses to Standardized Vocabulary Measures: Spanish

	EOWPVT Raw Within Language	EOWPVT Raw Single Language	BESAME Semantics	Lexical Revisions	Filled Pauses
EOWPVT Raw Within Language	1	.679**	.575**	.470**	.412*
EOWPVT Raw Single Language	.679**	1	.603**	.473**	.407*
BESAME Semantics	.575**	.603**	1	.345	.495**
Lexical Revisions	.470**	.473**	.345	1	.478**
Filled Pauses	.412*	.407*	.495**	.478**	1

Table 5.

Pearson Correlations of Lexical Revisions and Filled Pauses to Standardized Vocabulary Measures: English

	EOWPVT Raw Within Language	EOWPVT Raw Single Language	BESAME Semantics	Lexical Revisions	Filled Pauses
EOWPVT Raw Within Language	1	.952**	.863**	.287	-.026
EOWPVT Raw Single Language	.952**	1	.864**	.319	.005
BESAME Semantics	.863**	.864**	1	.214	.139
Lexical Revisions	.287	.319	.214	1	.445*
Filled Pauses	-.026	.005	.139	.445*	1

MLU AND NDW: ASSOCIATIONS WITH LEXICAL REVISIONS AND FILLED PAUSES

The third question investigated whether MLUw or NDW would be correlated with higher use of either lexical revisions or filled pauses in each language. MLUw and NDW were significantly correlated to each other in both English (0.649) and Spanish (0.690). BESAME Semantics in Spanish and English and raw EOWPVT within-language scores in Spanish were significant in differentiating the three language groups per vocabulary ability as observed in a Scheffe's contrast analysis.

In Spanish, NDW and MLUw both significantly correlated to only BESAME Semantics. When MLUw and NDW compared to the vocabulary standardized test measures in English, significant correlations were found for all three vocabulary measures: BESAME Semantics, EOWPVT within-language, and EOWPVT single-language conceptual. NDW showed more significant associations to total mazes, lexical revisions, and filled pauses than MLUw in both English and Spanish.

NDW. In English, the NDW for the TD groups was 99.3, 68.2 for the LN group, and 51.0 for the LI group. The NDW was significantly correlated to total mazes (0.702), lexical revisions (0.661), and filled pauses (0.542) in English.

In Spanish, the NDW was 105.50 for TD, 78.900 for LN, and 66.8 for LI. NDW was significantly correlated to total mazes (0.728), lexical revisions (0.417), and filled pauses (0.722).

MLUw. In English, MLUw was 6.9770 for TD, 6.2980 for LN, and 4.883 for LI. MLUw was not significantly correlated to lexical revisions nor filled pauses in English; however, was moderately correlated to total mazes (0.438).

In Spanish, MLU was 7.4250 for TD, 6.8890 for LN, and 5.97 for LI. MLUw was moderately correlated to filled pauses (0.424) and total mazes (0.377) but was not significantly correlated to lexical revisions. Tables 6 illustrates the correlations of MLUw and NDW with standardized vocabulary measures, lexical revisions, filled pauses, and total mazes in Spanish while Table 7 illustrates the correlations occurring in English.

Table 6.

Correlations for MLUw and NDW: Spanish

	MLUw	NDW
MLUw	1	.690**
NDW	.690**	1
Total Mazes	.377*	.728**
Lexical Revisions	.224	.417*
Filled Pauses	.424*	.722**
BESAME Semantics	.530**	.587**
EOWPVT Raw Single Language conceptual	.257	.361
EOWPVT Raw Within Language	.215	.283

Table 7.

Correlations for MLUw and NDW: English

	MLUw	NDW
MLUw	1	.649**
NDW	.649**	1
Total Mazes	.438*	.702**
Lexical Revisions	.340	.661**
Filled Pauses	0.361	.542**
BESAME Semantics	.616**	.584**
EOWPVT Raw Single Language conceptual	.514**	.577**
EOWPVT Raw Within Language	.471**	.543**

LANGUAGE EXPERIENCE, WITHIN-LANGUAGE, AND CROSS-LANGUAGE CORRELATIONS

The fourth purpose of the study was to investigate whether differences in language dominance and exposure affect lexical revisions and filled pauses in each language. To measure language dominance and exposure, age of English (AoE) acquisition and combined input/output exposure was gathered.

The LI group had less exposure to English overall (36.289%) compared to LN (51.8150%) and TD (40.059%), corresponding to the LI group having the most exposure to Spanish. Also, the LI group had later reported age of English acquisition, per parent

report. That is, whereas the TD group AoE acquisition was 3.278 and the LN group was 2.1, the LI group AoE acquisition was later at age 4.5.

Language dominance and exposure were not significantly correlated to lexical revisions or filled pauses in either Spanish or English. However, during the course of the analysis, within-language and cross-language patterns were found for lexical revisions and filled pauses.

Within-language. In English, lexical revisions and filled pauses demonstrated significant moderate within-language correlations of 0.445. In Spanish, lexical revisions and filled pauses also demonstrated significant moderate within-language correlations of 0.478.

Cross-language. Filled pauses in Spanish and in English were found to be significantly correlated (0.503). Lexical revisions in Spanish and English were moderately correlated (0.415). These findings suggest that the children who produce lexical revisions and filled pauses in one language are likely to produce them in the other language. Table 8 illustrates the within-language and cross-language correlations for lexical revisions and filled pauses.

Table 8.

Within-Language and Cross-Language Correlations for Lexical Revisions and Filled Pauses

	Age of English Acquisition	English Combined Input/ Output	English Lexical Revisions	English Filled Pauses	Spanish Lexical Revisions	Spanish Filled Pauses
Age of English Acquisition	1	-.655**	-.149	.147	-.147	.097
English Combined Input/ Output	-.655**	1	.022	-.241	.001	-.197
English Lexical Revisions	-.149	.022	1	.445*	.415*	.288
English Filled Pauses	.147	-.241	.445*	1	.327	.503**
Spanish Lexical Revisions	-.147	.001	.415*	.327	1	.478**
Spanish Filled Pauses	.097	-.197	.288	.503**	.478**	1

DISCUSSION

Maze production is multifactorial. Evidence suggests that different types of mazes are caused by underlying language and cognitive processes such as vocabulary level, language proficiency, working memory, and attention. Individuals with larger vocabularies may have more uncertainty in word selection, thereby producing more revisions and filled pauses (Bedore et al., 2006). MLU has been noted to be higher in utterances containing mazes (Miller, 1987; MacLachlan & Chapman, 1988; and Netterbladt & Hansson, 1999). Navarro-Ruiz and Rallo-Fabra (2001) reported that children with LI make less use of self-repairs and have a greater number of non-detected errors. Language proficiency has also been noted to affect rate of revisions. For example, Kaur, Hegde, Kumaraswamy, and Rao (2011) found that children less fluent in English, with more limited knowledge of the rules of the language, demonstrated less ability to make revisions.

Mazes are often referred to collectively as a group, yet they can be partitioned into categories including filled pauses, lexical revisions, repetitions, and grammatical revisions, each of which may reflect a distinctive role in language formulation. We specifically looked at filled pauses and lexical revisions to examine the role of vocabulary knowledge and language experience in their production. By examining lexical revisions and filled pauses, our intents were to extend existing research which conflicts regarding which types of mazes are produced by children of differing language ability as well as examine the theoretical implications of use of lexical revisions and filled pauses.

The current study found that the TD group exhibited significantly more lexical revisions in Spanish and significantly more filled pauses in English and Spanish compared to the LN and LI groups. As distinct maze behaviors, lexical revisions and filled pauses represent distinctive underlying language and cognitive processes. Lexical revisions indicate that the speaker has the metalinguistic capacity to recognize and repair specific linguistic breakdowns. In contrast, filled pauses are associated with linguistic uncertainty which does not directly correspond to recognizing or repairing linguistic breakdowns.

VOCABULARY KNOWLEDGE

To determine whether vocabulary knowledge was a factor in the production of lexical revisions or filled pauses, we analyzed several measures of vocabulary knowledge in relation to lexical revisions and filled pauses, including the BESAME Semantics scores, EOWPVT within-language and single-language conceptual scores, MLU, and NDW. We found that BESAME Semantics scores in English and Spanish and EOWPVT within-language scores in Spanish significantly differentiated the three language groups by language ability. The BESAME Semantics scores significantly correlated to filled pauses in Spanish, while EOWPVT within-language scores significantly correlated to lexical revisions; however, lexical revisions and filled pauses in English did not correlate to any standardized vocabulary measure.

When correlated with MLU and NDW, the results agree with findings by Bedore et al. (2006) which found that mazes in narratives of bilingual children aged 4-6 years

correlated with these language productivity measures. Bedore et al. (2006) found that in Spanish, grammatical revisions correlated with MLU and NDW while in English, filled pauses correlated with NDW and connectors were associated with MLU. Our study found that overall, NDW in narrative re-tell task was the best indicator for maze use amongst all measures analyzed in the study. NDW significantly correlated with total maze use, lexical revisions, and filled pauses in both English and Spanish. Interestingly, MLU did not demonstrate the same trend even though NDW and MLU were significantly correlated with each other. NDW may have been a better measure of expressive vocabulary ability than MLU since MLU is influenced by both morphosyntax and expressive vocabulary. If we had examined grammatical revisions in addition to lexical revisions, a correlation with MLU would likely have shown a greater effect. Dethorne, Johnson, and Loeb (2005) investigated MLU as a diagnostic measure and found that although semantics and morphosyntax interact and promote one another in language learning, MLU may confound these two relatively distinct domains. Therefore, the mazes produced by TD, LN, and LI groups may have been more directly related to vocabulary knowledge best represented by NDW than to the combination of vocabulary and morphosyntax that is calculated by MLU.

The inconsistency between standardized and language sample measures in regards to correlating with lexical revisions and filled pauses may be related to the differences in word recall for each of these tasks. The context for manipulating word retrieval during narratives may be more challenging than word retrieval during standardized testing due to new demands requiring planning, revising, and monitoring the evolving narrative. LN

and LI groups may also have decreased cognitive function for language, which would lead to a decreased ability to plan and revise, qualities which contribute to revisions.

An opposing argument states that maze production can also be caused by lower levels of vocabulary knowledge which in turn cause increased uncertainty. Word retrieval problems during the narrative tasks may be reflected by use of abandoned utterances or filled pauses. A typical behavior of children with word-finding problems is maze production, which children with LI have been noted to have (German, 1994). Children with LI have been noted to have word-finding problems. For instance, Fiestas, Peña, Bedore, and Sheng (2011) found that Spanish-English bilingual children with LI had difficulties describing functions of object nouns and providing category labels. LI children use nonspecific vocabulary, provide infrequent elaborations, and are unable to account for multiple or colloquial word meanings compared to children without LI (Gutiérrez-Clellen & DeCurtis, 1999).

Thordardottir and Ellis Weismer (2002) found that children with LI produced significantly more content mazes but fewer filled pauses than children with TD language during narration. Our study agreed that children with LI produced significantly fewer filled pauses in English and Spanish; however, we found that children with LI produced significantly less lexical revisions in Spanish as well, rather than more. In contrast, Loban (1976) found that individuals with low language proficiency exhibited more maze behavior and that the maze behavior remained consistent during the longitudinal study of thirteen years; however, the types of mazes produced were not disseminated across the different groups and the language group with low proficiency may certainly have

exhibited certain types of mazes not examined in our study that increased their maze count.

LANGUAGE EXPERIENCE, WITHIN-LANGUAGE, AND CROSS-LANGUAGE CORRELATIONS

We found that lexical revisions and filled pauses were generally observed most often in English than in Spanish except for the TD groups. We did not find any significant correlations between either language dominance or language exposure to either lexical revisions or filled pauses; rather, we found that lexical revisions and filled pauses presented with a stronger relationship to vocabulary knowledge than language dominance or exposure.

Regarding within-language effects on lexical revisions and filled pauses, we did find that lexical revisions and filled pauses were significantly correlated to each other within both English and Spanish languages. We also found that lexical revisions and filled pauses were significantly correlated across languages. Filled pauses were significantly correlated to each other in English and Spanish and lexical revisions in English and Spanish were moderately correlated to each other as well. These findings suggest that both lexical revisions and filled pauses may remain constant per individual regardless of language used, which indicates that linguistic uncertainty in only one language would not play a role.

CONCLUSION

The purpose of the study was to examine how vocabulary knowledge and language experience interact with lexical revisions and filled pauses in bilingual children with varying language abilities. Lexical revisions and filled pauses exhibited a stronger relationship with vocabulary knowledge than with language dominance or exposure. NDW proved to be the vocabulary measure that most significantly correlated with lexical revisions and filled pauses in both English and Spanish. If lexical revisions and filled pauses are assumed to represent the internal struggle to retrieve or select words, then the current study indicates that having a more extensive vocabulary knowledge increases the uncertainty in choosing words. A more extensive vocabulary knowledge would also explain why lexical revisions and filled pauses resulted in correlating with each other because the ability to revise requires the ability to recognize and repair linguistic breakdowns and the presence of filled pauses indicates uncertainty with word choice. The study agrees with findings by Bedore et al. (2006) that language productivity measures correlate with maze production. The study also partly converged with findings by Thordardottir and Ellis Weismer (2002) that filled pauses occurred more in typically-developing children, however disagreed that content mazes occurred more in language-impaired children.

Future research could focus on examining lexical revisions and filled pauses in bilingual compared to monolingual children. Bilingual children have equivalent but distributed vocabulary knowledge; however, they also display more instances of language disfluencies than monolingual children (Peña et al., 2002, Bedore et al., 2006). Future

research could also focus on continuing to discern the different types of mazes to further understand how language mirrors internal cognitive processes. The clinical implications of the study specify that use of lexical revisions and filled pauses are not an indication of decreased vocabulary knowledge; rather, lexical revisions and filled pauses are more likely indicators of vocabulary knowledge sufficient for revision and word choice.

References

Bedore, L. M., & Peña, E. D. (2008). Assessment of Bilingual Children for Identification of Language Impairment: Current Findings and Implications for Practice. *International Journal of Bilingual Education and Bilingualism*, *11*(1), 1–29.

Bedore, L.M., Fiestas, C.E., Peña, E.D, & Nagy, V.J. (2006). Cross-language comparisons of maze use in Spanish and English in functionally monolingual and bilingual children. *Bilingualism: Language and Cognition*, 9(3), 233-247.

Bedore, L. M., Peña, E. D., García, M., & Cortez, C. (2005). Clinical forum: Conceptual versus monolingual scoring: When does it make a difference?. *Language, Speech, and Hearing Services in Schools*, *36*(3), 188-200.

Bedore, L. M., & Leonard, L. B. (2001). Grammatical morphology deficits in Spanish-speaking children with specific language impairment. *Journal of Speech, Language and Hearing Research*, *44*(4), 905.

Bracken, B. A., & McCallum, R. S. (1998). *Universal Nonverbal Intelligence Test*

de Bot, K. (1996). The psycholinguistics of the output hypothesis. *Language Learning*, *46*, 529–555.

DeJoy, D.A. and Gregory, H. (1985). The relationship between age and frequency of disfluency in preschool children. *Journal of Fluency Disorders, 10*(2), 107-122.

Dethorne, L.S., Johnson, B.W., & Loebe, J.W. (2005). A closer look at MLU: What does it really measure? *Clinical Linguistics & Phonetics, 19*(8), 635-648.

Dollaghan, C. and Campbell, T. (1992). A procedure for classifying disruptions in spontaneous language samples. *Topics in Language Disorders*, *12*, 56-68.

Fiestas, C., Peña, E. D., Bedore, L. M., & Sheng, L. (2011, November). The definition skills of bilingual children with language impairment. *Poster presented at the annual convention of the American Speech-Language-Hearing Association*, Philadelphia, PA.

Fletcher, P. (1991). Evidence from syntax for language impairment. In J. Miller (Ed.), *Research on Child Language Disorders: a decade of progress* (Austin, TX: Pro-Ed.), pp. 169-187.

German, D. (1994). Word Finding difficulties in children and adolescents. In G. Wallach and K. Butler (Eds), *Language Learning Disabilities in School-Age Children and Adolescents. Some principles and applications* (New York: Maxwell Macmillan International), pp. 323-347.

Gollan,T.H.& Silverberg,N.B. (2001).Tip-of-the-tongue states in Hebrew–English bilinguals. *Bilingualism: Language and Cognition*, *4*(1), 63–84.

Gollan, T. H. & Acenas, L. A. (2004). What is a TOT? Cognate and translation effects on tip-of-the-tongue states in Spanish–English and Tagalog–English bilinguals. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *30*(1), 246–269.

Gollan, T. H., Montoya, R. I. & Bonanni, M. P. (2005). Proper names get stuck on bilingual and monolingual speakers' tip of the tongue equally often. *Neuropsychology*, 19 (3), 278–287.

Gollan, T.H., Montoya, R.I., Fennema-Notestine, C., and Morris, S.K. (2005). Bilingualism affects picture naming but not picture classification. *Memory and Cognition, 33,* 1220-1234.

Gollan, T.H., Montoya, R.I., Cera, C., and Sandoval, T.C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language, 58,* 787-814.

Gollan, T.H., Slattery, T.J., Goldenberg, D., van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. *Journal of Experimental Psychology: General, 140,* 186-209.

Gollan, T.H., Sandoval, T., & Salmon, D.P. (2011). Cross-language intrusion errors in aging bilinguals reveal the link between executive control and language selection. *Psychological Science, 22,* 1155-1164.

Gordon, P.A., Luper H.L., and Peterson, H.A. (1986).The effects of syntactic complexity on the occurrence of disfluencies in 5 year old nonstutterers. *Journal of Fluency Disorders, 11,* 151-164.

Green, D. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition, 1,* 67-81.

Guo, T., Liu, H., Misra, M.. and Kroll, J.F. (2011). Local and global inhibition in bilingual word production: fMRI evidence from Chinese-English bilinguals. *NeuroImage, 56,* 2300-2309.

Gutiérrez-Clellen, V. F., & Kreiter, J. (2003). Understanding child bilingual acquisition using parent and teacher reports. *Applied Psycholinguistics, 2,* 267–288.

Gutiérrez-Clellen, V. F., & DeCurtis, L. (1999). Word definition skills in Spanish-speaking children with language impairment. *Communication Disorders Quarterly, 21,* 23-31.

Harrington, M., & Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition*, *14*(01), 25–38.

Holowka, S., Brosseau-Lapre´, F. and Petitto, L.A. (2002). Semantic and conceptual knowledge underlying bilingual babies' first signs and words. *Language Learning, 52,* 205-262.

Hoshino, N., & Thierry, G. (2011). Language selection in bilingual word production: Electrophysiological evidence for cross-language competition, *Brain Research, 1371,* 100-109.

Kaur, R., Hegde, M., T.S.S., Kumaraswamy, & Rao (2011). Mazes in typically developing bilingual children. *Asia Pacific Journal of Speech, Language and Hearing*, *14*(4), 197-203.

Kormos, J. (1999). Monitoring and self-repair in L2. *Language Learning*, *49*(2), 303–342

Kroll, J. F., & Gollan, T. H. (2014). Speech planning in two languages: What bilinguals tell us about language production. In V. Ferreira, M. Goldrick, & M. Miozzo (Eds.). *The Oxford handbook of language production*. Oxford: Oxford University Press.

Leadholm, B. J. & Miller, J. (1995). *Language sample analysis:The Wisconsin guid*e. Madison, WI: Wisconsin Department of Public Health.

Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning*, 40 (3), 387–417.

Levelt, W. (1989). *Speaking: from intention to articulation* (Cambridge, MA: MIT Press).

Levelt, W., Roelofs, A. & Meyer, A. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*(1), 1–75.

Loban, W. (1976). *Language development: Kindergarten through grade twelve*. Urbana, IL: National Council of Teachers of English.

Nettelbladt, U. & Hansson, K. (1999). Mazes in Swedish pre-school children with specific language impairment. *Clinical Linguistics and Phonetics, 13*(6), 483–497.

MacLachlan, B. and Chapman, R. (1988). Communication breakdowns in normal and language learning-disabled children's conversation and narration. *Journal of Speech and Hearing Disorders, 53*, 2-7.

Marchman, V.A., Martınez-Sussmann, C. and Dale, P.S. (2004). The language-specific nature of grammatical development: Evidence from bilingual language learners. *Developmental Science, 7*, 212-224.

Miller, J. and Klee, T. (1995). Computational approaches to the analysis of language impairment. In P. Fletcher and B. MacWhinney (Eds), *The Handbook of Child Language* (Oxford: Basil Blackwell), pp. 545-572.

Miller, J. (1987). A grammatical characterisation of language disorder. In J. Martin, P. Fletcher, P. Grunwell and D. Hall (Eds), *Proceedings of the First International Symposium on Specific Speech and Language Disorders in Children* (London: AFASIC), pp. 100-113.

Misra, M., Guo, T., Bobb, S.C., and Kroll, J.F. (2012) When bilinguals choose a single word to speak: Electrophysiological evidence for global inhibition in bilingual word production. *Journal of Memory and Language, 67,* 224-237.

Navarro-Ruiz, M. I. & Rallo-Fabra, L. (2001). Characteristics of mazes produced by SLI children. *Clinical Linguistics and Phonetics*, *15*(1 & 2), 63–66.

Paradis, J., Crago, M., Genesee, F. and Rice, M.L. (2003). French-English bilingual children with SLI: How do they compare with their monolingual peers? *Journal of Speech, Language and Hearing Research, 46*, 113-127.

Patterson, J.L. (2000). Observed and reported expressive vocabulary and word combinations in bilingual toddlers. *Journal of Speech, Language and Hearing Research, 43*(1), 121-128.

Patterson, J.L. (1998). Expressive vocabulary development and word combinations of Spanish-English bilingual toddlers. *American Journal of Speech Language Pathology. 7*(4), 46-56.

Pearson, B.Z., Fernandez, S.C. and Oller, D.K. (1995). Cross-language synonyms in the lexicons of bilingual infants: One language or two? *Journal of Child Language, 22*(2), 345-368.

Pearson, B.Z. and Fernandez, S.C. (1994). Patterns of interaction in the lexical growth in two languages of bilingual infants and toddlers. *Language Learning, 44*(4), 617-653.

Pearson, B.Z., Fernandez, S. and Oller, D.K. (1993) Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. *Language and Learning, 43*(1), 93-120.

Peña, E.D., Bedore, L.M. and Zlatic-Giunta, R. (2002). Category generation performance of young bilingual children: The influence of condition, category, and language. *Journal of Speech, Language, and Hearing Research, 41*, 938-947.

Peña, E. D., Iglesias, A., & Lidz, C. S. (2001). Reducing test bias through dynamic assessment of children's word learning ability. *American Journal of Speech-Language Pathology, 10,* 138-154.

Poulisse, N., & Bongaerts, T. (1994). First language use in second language production. *Applied Linguistics*, *15*, 36–57.

Poulisse, N. (1997). Language production in bilinguals. In A. deGroot & J. Kroll (eds.). *Tutorials in bilingualism: Psycholinguistic perspectives*, pp. 201–224. Mahwah, NJ: Lawrence Erlbaum Associates.

Poulisse,N. (1999). *Slips of the tongue: Speech errors in first and second language production.* Amsterdam: John Benjamins.

Rieger, C. L. (2003). Repetitions as self-repair in English and German conversations. *Journal of Pragmatics*, 35 (1), 47–69.

Robinson, P. (1995). Attention, memory and the "noticing" hypothesis. *Language Learning*, *45*, 283–331.

Sheng, L., Pena, E. D., Bedore, L. M., & Fiestas, C. E. (2012). Semantic deficits in Spanish-English bilingual children with language impairment. *Journal of Speech, Language, and Hearing Research*, *55*(1).

Taliancich-Klinger, C. L., Byrd, C. T., & Bedore, L. M. (2013). The disfluent speech of a Spanish–English bilingual child who stutters. *Clinical Linguistics & Phonetics*, *27*(12), 888–904.

Thordardottir, E. & Ellis Weismer, S. (2002). Content mazes and filled pauses on narrative language samples of children with specific language impairment. *Brain and Cognition*, *48*(2–3), 587–592.

van Hest, E. (1996b). *Self-repair in L1 and L2 production*. Tilburg, The Netherlands: Tilburg University Press.

Verhoeven, L.T. (1989). Monitoring in children's second language speech. *Second Language Research*, *5*(2), 141-155.

Wiese, R. (1984). Language production in foreign and native languages: Same or different? In H. Dechert, D. Möhle & M. Raupach (eds.), *Second language productions*, pp. 11–25. Tübingen: Narr.